

文章编号:1672-9331(2020)02-0061-07

基于文本挖掘的高处坠落事故致因及关联规则分析

李 珏, 李世杰

(长沙理工大学 交通运输工程学院, 湖南 长沙 410114)

摘 要: 高处坠落事故是建筑业中最常见的事故之一。为明确高处坠落事故致因,找到有效的事故预防措施,搜集了 2012~2017 年国内 528 例高处坠落事故调查报告,以 R 语言为平台,结合文本挖掘技术确定了事故中的 32 项事故致因,其中包括 7 项关键事故致因。利用 Apriori 算法挖掘出事致因之间的关联规则,通过可视化关联规则结果,得到了高处坠落事故发生过程中的事故致因因果关系,为高处坠落事故的预控提供参考。

关键词: 文本挖掘;高处坠落事故;事故致因;R 语言;TF-IDF 算法;关联规则;Apriori 算法

中图分类号: X947

文献标识码: A

Analysis of causes and association rules for falling accidents based on text mining method

LI Jue, LI Shi-jie

(School of Traffic and Transportation Engineering, Changsha University of Science &
Technology, Changsha 410114, China)

Abstract: Falling accidents is one of the most common accidents in construction industry. In order to clarify causes of falling accidents and find effective prevention measures, 528 domestic investigation reports of falling accidents from 2012 to 2017 were collected. Based on R language and text mining technology, 32 causes of accident were identified, including 7 key causes. Apriori algorithm was used to mine association rules among the accident causes. By visualizing the results of the association rules, the causalities of the accident causes during the falling accidents were obtained, which could provide reference for the prevention of falling accidents.

Key words: text mining; falling accident; accident cause; R language; TF-IDF algorithm; association rule; Apriori algorithm

高处坠落事故具有发生频繁、危险程度高、伤亡大等特点,是建筑行业中最常见的安全事故之一^[1]。在近几年的房屋市政工程生产安全事故中,“高处坠落、物体打击、起重伤害、坍塌”这四大

伤害造成的事故累计占到了建筑安全事故总数的近 90%^[2],其中高处坠落事故的发生率居于首位。分析和找出引起高处坠落事故的原因,以及各因素之间的关系,并有针对性的采取防治措施,是预

收稿日期:2019-12-11

基金项目:国家自然科学基金资助项目(71371036);湖南省研究生科研创新项目(CX20190669))

通讯作者:李 珏(1978-),男,湖南平江人,长沙理工大学副教授,主要从事工程项目管理等方面的研究。

E-mail:lij93@csust.edu.cn

防高处坠落事故发生的重要环节。

国内许多学者针对高处坠落事故进行了相关的研究工作。邓航等^[3]运用事故树法分析高处坠落事故发生的基本原因,从技术和管理方面提出预防对策;易灿南等^[4]将事故树分析法与层次分析法相结合,发现高处坠落事故中坠前防护对事故的影响最大;郑霞忠等^[5]建立了高处坠落人因失误结构方程模型,对人为失误因素进行路径分析并提出了预防人因干预的策略;谭钦文等^[6]提出了事故树“人-机-环”三因素展开模型,用于高处坠落事故分析,得到了导致事故发生的主要影响因素。

回顾高处坠落事故的研究成果可以发现,国内学者所使用的研究方法,在识别事故致因时通常需要阅读事故相关文件或进行个案分析,容易受到主观判断的影响。因此,本研究在前人研究的基础上,为减少人工进行事故致因识别的主观性和工作量,提出使用文本挖掘技术对高处坠落事故致因进行分析,为高处坠落事故致因分析提供一种新的思路。

1 文本挖掘技术

文本挖掘(text mining)是一种从非结构化文本信息中获取有价值的知识和信息的计算机处理技术,又可称作知识发现技术,可应用到统计学、自然语言学、机器学习等多个领域,为各类事物的研究提供新的方法^[7]。文本挖掘的基本步骤一般为:文本数据收集、文本预处理、数据结构化、数据分析、结果可视化、知识发现。

近年来,文本挖掘技术得到广泛运用,Kim等^[8]通过文本挖掘技术对苹果公司的专利进行分析,揭示了苹果公司技术发展的趋势和各技术之间的联系。在火灾事故的研究中,文本挖掘与主成分分析法相结合,推断出施工区域火灾事故的季节性因素^[9]。Jie等^[10]将文本挖掘方法应用于安全风险分析中,从156份事故报告中确定了15个安全风险因素和3类参与者;吴伋等^[11]使用文本挖掘方法,分析大量的内河船舶碰撞事故报告,识别出了事故致因,并验证了其结果的准确性;谭章禄等^[12]借助文本挖掘,研究煤矿安全隐患信息的关联和分布,证明了将文本挖掘用于安全隐患

分析具有可行性。

1.1 文本挖掘工具选取

本研究选用R语言对高处坠落事故报告进行文本挖掘。R语言是一个自由、免费且源代码开放的平台,可以用于统计计算与统计制图,具有一套完整的数据处理、计算和绘图系统。R语言在数据挖掘方面具有较强的优势^[13],自带了许多能用于各个学科的函数程序包,在使用R语言的过程中调用这些函数,就可以构造出各个学科、领域所需要的数学模型。R语言有众多的算法程序包,用户可以根据自己的需求进行选择,从而达到在文本挖掘过程中对数据进行统计分析、数据处理和可视化的目的。

1.2 文本挖掘语料选取

事故调查报告是安全风险事故分析的重要数据来源。事故调查报告内容包括事故相关单位及工程概况,事故发生时间、地点、现场情况、简要经过、造成的伤亡人数,事故经济损失和人员伤亡情况,事故发生直接与间接原因,事故责任认定及处理建议等。与其它文件相比,事故调查报告对事故的描述更准确也更丰富,故本研究选择事故报告作为文本挖掘的语料来源。

本研究收集了我国2012~2017年528例高处坠落事故报告,涉及19个省级行政区(见表1)。事故报告来源于国家安监总局、国家住建部、各省市相关的行政管理部门,报告数量较为丰富,可以保证使用文本挖掘分析高处坠落事故致因的客观性。

表1 事故发生省级行政区域统计表

Table 1 Statistics of provinces with accidents

省级行政区	数量/例	省级行政区	数量/例
北京	28	福建	12
天津	14	江西	1
上海	35	山东	15
重庆	24	河南	6
山西	3	湖北	23
江苏	97	湖南	15
浙江	31	广东	129
安徽	7	四川	53
贵州	1	云南	8
广西	26		

2 基于文本挖掘的高处坠落事故致因

2.1 高处坠落事故致因挖掘

使用 R 语言对高处坠落事故报告进行预处理。为了使分词达到理想的效果,避免出现专业词汇无法识别以及虚词干扰的情况,在进行分词之前选用专业术语词库和自定义去虚词表来提高文本挖掘的精确度。专业词库为搜狗词库的专有名词词库及作者手动生成的词库,去虚词表根据《现代汉语虚词词典》制作而成。进行分词之后得到 308 项分词结果作为初始的特征项。

初始特征项中,“事故”“安全”“生产”出现频率为前三,类似表示文本性质的词汇的出现频率也非常高,这是因为文本挖掘的对象是高处坠落事故报告。因此,文本中上述词汇将会大量出现,但是对于研究高处坠落事故致因并无帮助,所以

删除。同样,也删除“监理”“施工人员”等具有较高词频的词。

在进行事故调查报告的预处理以及特征项筛选之后,得到了初步的事故致因特征项。由于事故调查报告非统一模板,即使是相同的事故致因也可能用不同的表述方式,导致分词结束后会出现许多具有相同意义的不同词汇,比如“安全管理不到位”“安全管理缺失”等词汇所表达的都是安全管理不到位的意思。将这些词都用同一词汇替换,这样在生成向量空间模型时可以将这些具有相同意义的词合并在同一维度,避免某些由于表述较为生僻而出现次数太少的词汇在分析时被忽略,影响了文本挖掘的准确性。相同意义词汇合并见表 2。

使用归并词群表合并同语义词汇并剔除多余高频词汇后,挖掘出 32 项具有分析价值的高频事故致因特征项,见表 3、图 1。

表 2 归并词群表
Table 2 Merged thesaurus

替换词	同义词
安全教育不到位	缺乏安全教育、安全教育不到位、安全教育培训不到位…
安全管理不到位	安全生产管理不到位、安全生产工作不到位…
安全意识淡薄	安全意识不强、安全意识淡薄、缺乏安全意识…
制度未落实或不健全	安全责任制未落实、安全生产责任制度未建立、安全管理制度未建立…
违规违章行为	违反高处作业操作规程、违法安装、违章作业…
无资质或无证施工	未取得《施工许可证》、无工程施工资质、无资质…
责任履行不到位	责任履行不到位、履行安全生产管理职责不到位…
监理不到位	未认真履行现场监理工作职责、安全监理职责不到位、监理不到位…
未系安全带	未系安全带、未佩戴安全带…
⋮	⋮

表 3 高处坠落事故主要事故致因特征项
Table 3 Feature items of main causes of falling accidents

代号	致因项	代号	致因项	代号	致因项
T ₁	安全监管	T ₁₂	技术交底	T ₂₃	防护门
T ₂	安全教育不到位	T ₁₃	未戴安全帽	T ₂₄	摔倒
T ₃	有安全隐患	T ₁₄	采光不好	T ₂₅	天气
T ₄	违规违章行为	T ₁₅	制度未落实或不健全	T ₂₆	未能认真履行
T ₅	未系安全带	T ₁₆	安全绳	T ₂₇	安全设备
T ₆	安全意识淡薄	T ₁₇	安全措施	T ₂₈	打滑
T ₇	防护措施不到位	T ₁₈	责任履行不到位	T ₂₉	不符合国家标准
T ₈	无资质或无证施工	T ₁₉	防护网	T ₃₀	滑倒
T ₉	安全管理不到位	T ₂₀	监理不到位	T ₃₁	安全生产责任未落实
T ₁₀	劳动防护用品	T ₂₁	防护栏	T ₃₂	异常气候
T ₁₁	安全检查	T ₂₂	管理混乱		

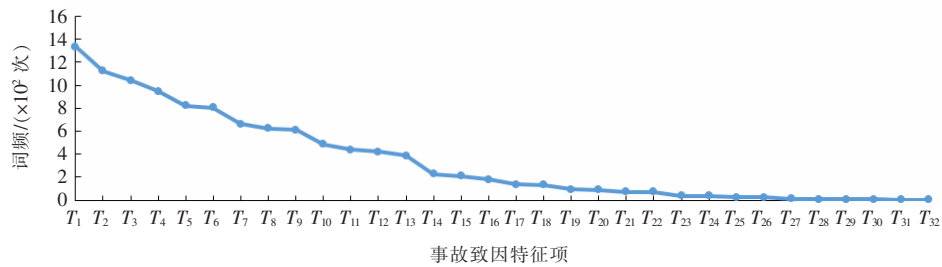


图1 高处坠落事故主要事故致因频次

Fig. 1 The frequency of main causes of falling accidents

2.2 事故致因空间向量模型构建

在文本挖掘的过程中,词频并不能完全反映一个词在整体中的重要性,为了找到事故致因中的关键项,需要将文本转化为向量空间模型^[14]。向量空间模型由特征项和特征项权重组成,通过2.1与2.2的处理,即可得到所需要的特征项。在向量空间模型中 D 表示一篇文档, T_k 表示挖掘出来的特征项。使用TF-IDF算法^[15]给特征项赋予权重,公式如下:

$$tf_{i,j} = \frac{n_{i,j}}{\sum_k n_{k,j}} \quad (1)$$

$$idf_i = \log \frac{|D|}{|\{j:t_i \in d_j\}|} \quad (2)$$

$$W_{ij} = tf_{i,j} \times idf_i \quad (3)$$

式中: $n_{i,j}$ 为特征项在文档 D_i 中出现的次数;

$\sum_k n_{k,j}$ 为所有文档 $|D|$ 中特征项出现的总次数; $|\{j:t_i \in d_j\}|$ 为含此特征项的文档数目; $|D|$ 为所有文档的数量。由公式(3)即可求出每个特征项在文本中的权重 W_{ij} 。

在R语言中调用tm包构建向量空间模型,如表4所示。其中, T_i 为高处坠落事故的第 i 项事故致因; D_i 为第 i 份高处坠落事故调查报告。表4中列举了部分文本和特征项的空间向量模型。

2.3 高处坠落事故致因词云展示

为了能够更加直观的展示所挖掘出来的事故致因重要性程度,对空间向量模型的数据进行可视化,得到图2所示高处坠落事故致因词云。图2中词语字体的大小与位置代表该词语的重要程度,字体由大至小、位置由内至外代表重要程度在递减。

表4 高处坠落事故调查报告空间向量模型

Table 4 Spatial vector model of falling accident investigation reports

D_i	T_i										
	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}	T_{11}
D_1	0	0.1	0.8	0	0	0.2	0	0	0	0	0
D_2	0	0	0	0.1	0.3	0.1	0	0	0	0	0
D_3	0	0	0	0.1	0	0	0	0.4	0	0	0
D_4	0	0.1	0	0.1	0.1	0	0	0	0	0.1	0
D_5	0	0	0	0	0	0	0	0	0	0	0
D_6	0	0	0	0	0	0	0	0	0.1	0	0
D_7	0	0	0	0	0	0	0.1	0.3	0	0.1	0
D_8	0	0	0	0.1	0	0	0	0	0	0	0
D_9	0	0	0	0.1	0	0	0	0	0	0	0
D_{10}	0	0.1	0	0.1	0	0	0.2	0	0	0	0
D_{11}	0	0	0	0	0	0.1	0.1	0.2	0.1	0	0
D_{12}	0	0	0	0	0.2	0	0.1	0	0	0.1	0

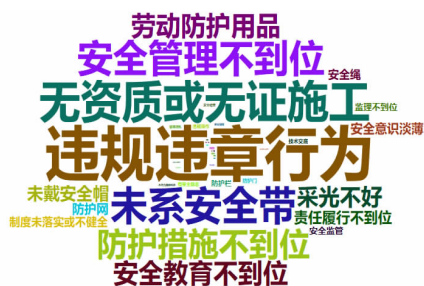


图 2 高处坠落事故致因词云

Fig. 2 Nephogram of falling accident causes

按照词云图中词汇的大小和排布位置,可以将其区分为关键致因与一般致因。其中关键致因包括:违规违章行为(0.133)、无资质或无证施工(0.1)、未系安全带(0.09)、安全管理不到位(0.09)、防护措施不到位(0.07)、劳动防护用品(0.06)、安全教育不到位(0.06)7项,其余则为一般致因。括号中所示数据为该致因在图中所占权重。

3 高处坠落事故致因关联规则分析

在以上的分析中,比较清晰地展现了高处坠落事故的致因。但是根据事故致因轨迹交叉理论^[16],事故的发生过程可以描述为事故致因导致的运动轨迹,具有复杂的因果关系。仅仅依靠单个事故致因无法制定有效的高坠事故防控手段。本研究将采用关联规则对事故致因进行分析,得到具有强关联规则的事故致因因果关系。

3.1 关联规则挖掘

关联规则是一种常见的数据挖掘技术,能够挖掘数据库中各项集之间的关系^[17]。支持度(Support)、置信度(C Confidence)、提升度(Lift)是关联规则中最重要的三个概念。支持度为事务集中项集 X 与 $Y(X \cup Y)$ 和总项集数之比,即出现的概率;置信度为项集 $X \cup Y$ 同时出现的次数与项集 X 出现的次数之比,用 $\text{Confidence}(X \rightarrow Y)$ 表示,即在项集 X 出现的条件下,又出现项集 Y 的概率;提升度表示为 $\text{Lift}(X \rightarrow Y)$,为 $X \rightarrow Y$ 的置信度与项集 Y 的支持度之比,意义在于度量项集 X 和项集 Y 的独立性,提升度大于 1,则强关联规则有效^[18]。

本研究采用 Apriori 算法^[19]挖掘事故致因之

间的关联。该算法包括两大部分:1) 搜索频繁集;2) 依据频繁项集产生关联规则。在 R 语言中调用 arules 包进行关联规则分析。

根据关联规则所需要的数据格式,利用事故调查中包含的特征项构建布尔矩阵,如表 5 所示。每列代表一个 item,即为文本挖掘出来的事故致因项;每行代表一个 transaction,即事故调查报告 D_i ;数字 1 表示该事故调查报告中出现了此事故致因;0 表示没有出现。搜索频繁集,然后设置最小置信度为 0.6,并将提升度大于 1 的规则筛选出来,阈值参照郭圣煜等^[20]的研究并经过试验后选取,由此得到 53 条关联规则。表 6 展示了部分事故致因关联规则。

表 5 关联规则挖掘布尔矩阵

Table 5 Boolean matrix used by association rules

	安全教育 不到位	有安全 隐患	...	监理 不到位	未戴 安全帽
D_1	1	1	...	0	0
D_2	0	0	...	0	1
D_3	1	1	...	0	0
D_4	1	1	...	1	0
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots

表 6 部分事故致因关联规则

Table 6 Part of association rules for accident causes

规则	支持度	置信度	提升度
安全意识淡薄 \Rightarrow 安全教育不到位	0.515	0.749	1.061
有安全隐患 \Rightarrow 安全教育不到位	0.494	0.765	1.083
违规违章行为 \Rightarrow 安全意识淡薄	0.403	0.717	1.043
\vdots	\vdots	\vdots	\vdots
安全绳 \Rightarrow 安全教育不到位	0.100	0.828	1.172

3.2 关联规则可视化

利用 R 语言中的 arulesViz 包将所得到的 53 条关联规则进行可视化处理,得到如图 3 所示的高处坠落事故致因关联规则关系图。顶点代表项集,边表示规则中关系的关联规则,两项集之间的圆点代表关联规则的支持度与提升度,圆点的尺寸越大、颜色越深则该关联规则的支持度越高、提升度越大。

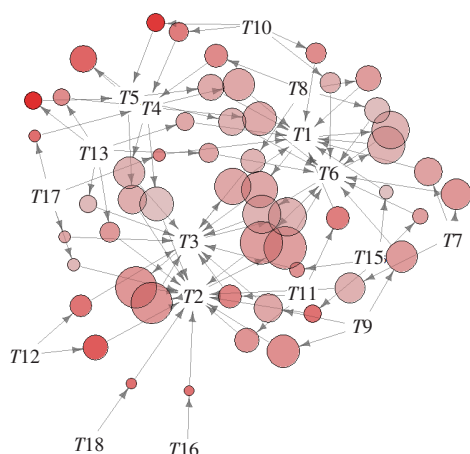


图3 高处坠落事故致因关联规则关系图

Fig. 3 Association rule diagram of falling accident causes

3.3 综合分析

根据图3的结果可以看出各致因之间的因果关系。安全教育不到位 T_2 、安全意识淡薄 T_6 、安全监管 T_1 、有安全隐患 T_3 处于整个关系图的中心位置,有13条边指向安全教育不到位(支持度为0.100~0.515),11条边指向有安全隐患(支持度为0.123~0.494),11条边指向安全意识淡薄(支持度为0.170~0.515),11条边指向安全监管(支持度为0.131~0.451)。

说明上述4项致因与其它致因联系紧密,并且具有较高的支持度,是引起其它事故致因的根源,始终贯穿在所有的高处坠落事故中。根据海因里希提出的事故多米诺理论,并结合图3的关联规则关系可视化,可知当这4项处于源头的致因存在时,极有可能导致在后续的施工过程中出现物的不安全状态、人的不安全行为以及管理的混乱与不到位。比如图中与之有强关联的防护措施不到位 T_7 、出现违规违章行为 T_4 和安全管理不到位 T_9 等致因,而由这些致因又会导致下一个致因的出现,形成多米诺效应,最终导致高处坠落事故发生。因此,针对这4项致因采取相应的改进措施,能够最大限度的减少后续事故致因的出现,保证高处坠落事故可控可防。

结合词云分析结果,违规违章行为、无资质或无证施工、未系安全带、安全管理不到位、防护措施不到位、劳动防护用品、安全教育不到位7种关键事故致因,可以看作是整事故致因关联规则

因果关系图中的重要节点。对这些节点及时进行干涉,能中断事故的发展过程,避免事故多米诺效应的继续,从而有效减少高处坠落事故的发生。

安全教育不到位既是7种关键事故致因之一,又是关联规则分析中的根源致因,在高处坠落事故预防与控制中占有极其重要的地位。由于建筑行业中采用劳务分包十分频繁,从业人员在上岗之前缺少安全知识教育,导致安全意识始终处在一个较低的水平。成功的安全教育是保障施工安全的前提,只有加大安全教育力度,将安全教育落到实处,才能提高工程中各方人员的安全意识,减少不安全现象的发生。

4 结论

1) 利用文本挖掘技术和R语言,通过分词、特征项选取、使用TF-IDF算法构建空间向量模型、词云可视化。从528份非结构化的事故调查报告中确定了32项高处坠落事故致因,其中违规违章行为(0.133)、无资质或无证施工(0.1)、未系安全带(0.09)、安全管理不到位(0.09)、防护措施不到位(0.07)、劳动防护用品(0.06)、安全教育不到位(0.06)7项权重较高,作为关键致因,其余25项为一般致因。

2) 采用Apriori算法,挖掘高处坠落事故致因间的关联规则,并将其因果关系可视化。发现安全教育不到位覆盖了13条规则,支持度为0.100~0.515;安全意识淡薄覆盖了11条规则,支持度为0.170~0.515;有安全隐患覆盖了11条规则,支持度为0.123~0.494;安全监管覆盖了11条规则,支持度为0.131~0.451。此4项致因与其它致因联系紧密,是引起其它事故致因的根源。结合词云与关联规则分析,安全教育在高处坠落事故预防工作中具有重要的基础性地位。在建筑工程中,良好的安全教育能够减少后续事故致因的出现,而在高处施工中,对7项关键事故致因进行重点防控,能够中断事故发展过程,有效预防事故发生。

3) 本研究的局限性在于各事故调查报告内的叙述表达方式不同和文本挖掘方法对句子的理解错误,都可能使得某些事故致因被遗漏。在后续

研究中可对文本挖掘算法进行进一步地优化,增强文本挖掘方法的语义分析能力,从非结构化的文本数据中找到更多有用信息。

〔参考文献〕

- [1] 张明轩,朱月娇,翟玉杰,等. 建筑工程高处坠落事故的故障树分析方法研究[J]. 煤炭工程,2008(2): 112-114.
ZHANG Ming-xuan, ZHU Yue-jiao, ZHAI Yu-jie, et al. Study on fault tree analysis method for falling accidents at heights of construction[J]. Coal Engineering, 2008(2): 112-114.
- [2] 中华人民共和国住房和城乡建设部. 房屋市政工程施工安全事故情况通报(2014~2017)[EB/OL]. <http://www.mohurd.gov.cn>.
Ministry of Housing and Urban-Rural Construction of the People's Republic of China. Bulletin on production safety accidents of housing and municipal engineering(2014-2017)[EB/OL]. <http://www.mohurd.gov.cn>.
- [3] 邓航,庞奇志,朱德英,等. 建筑施工高处坠落事故分析及预防对策[J]. 工业安全与环保,2010,36(4): 57-59.
DENG Hang, PANG Qi-zhi, ZHU De-ying, et al. Analysis and countermeasures about falling accidents of architecture construction[J]. Industrial Safety and Environmental Protection, 2010, 36(4): 57-59.
- [4] 易灿南,胡鸿,廖可兵,等. FTA-AHP 方法研究及应用[J]. 中国安全生产科学技术,2013,9(11): 167-173.
YI Can-nan, HU Hong, LIAO Ke-bing, et al. Research on FTA-AHP and its application[J]. Journal of Safety Science and Technology, 2013, 9(11): 167-173.
- [5] 郑霞忠,王晓宇,陈述,等. 高处坠落事故的人因失误与干预策略研究[J]. 中国安全生产科学技术,2017,13(6):139-144.
ZHENG Xia-zhong, WANG Xiao-yu, CHEN Shu, et al. Study on human error and intervention strategies of high falling accidents[J]. Journal of Safety Science and Technology, 2017, 13(6): 139-144.
- [6] 谭钦文,董勇,段正肖,等. 复杂事件事故树“人-机-环”简易展开模型及应用[J]. 工业安全与环保,2018,44(5):57-60.
TAN Qin-wen, DONG Yong, DUAN Zheng-xiao, et al. Simple expansion model of “man-machine-environment” fault tree in complex system and its application[J]. Industrial Safety and Environmental Protection, 2018, 44(5): 57-60.
- [7] Fayyad U, Piatetsky-Shapiro G, Smyth P. The KDD process for extracting useful knowledge from volumes of data[J]. Communications of the ACM, 1996, 39(11): 27-34.
- [8] Kim J M, Jun S. Graphical causal inference and copula regression model for apple keywords by text mining[J]. Advanced Engineering Informatics, 2015, 29(4): 918-929.
- [9] Joon-Soo Kim, Byung-Soo Kim. Analysis of fire-accident factors using big-data analysis method for construction areas[J]. KSCE Journal of Civil Engineering, 2018, 22(5): 1 535-1 543.
- [10] Jie L, Jian-ping W, Na X, et al. Importance degree research of safety risk management processes of urban rail transit based on text mining method[J]. Information, 2018, 9(2): 26.
- [11] 吴俊,江福才,姚厚杰,等. 基于文本挖掘的内河船舶碰撞事故致因因素分析与风险预测[J]. 交通信息与安全,2018,36(3): 8-18.
WU Ji, JIANG Fu-cai, YAO Hou-jie, et al. An analysis and risk forecasting of inland ship collision based on text mining[J]. Journal of Transport Information and Safety, 2018, 36(3): 8-18.
- [12] 谭章禄,陈晓,宋庆正,等. 基于文本挖掘的煤矿安全隐患分析[J]. 安全与环境学报,2017,17(4): 1 262-1 266.
TAN Zhang-lu, CHEN Xiao, SONG Qing-zheng, et al. Analysis for the potential hazardous risks of the coal mines based on the so-called text mining[J]. Journal of Safety and Environment, 2017, 17(4): 1 262-1 266.
- [13] 朱昶胜,王莎莎,王永贤. 基于 R+Hadoop 的中药材大数据的分析及预测[J]. 兰州理工大学学报,2017,43(1): 98-103.
ZHU Chang-sheng, WANG Sha-sha, WANG Yong-xian. Analysis and prediction of big data of Chinese medicinal materials based on R+Hadoop[J]. Journal of Lanzhou University of Technology, 2017, 43(1): 98-103.

(下转至第 74 页)

- 102-106.
- [8] 赵敏. 基于遗传算法优化神经网络的再生保温混凝土强度预测[D]. 太原: 太原理工大学, 2018.
- ZHAO Min. Strength prediction of recycled insulated concrete based on genetic algorithm optimization neural network[D]. Taiyuan: Taiyuan University of Technology, 2018.
- [9] 韩小孩, 张耀辉, 孙福军, 等. 基于主成分分析的指标权重确定方法[J]. 四川兵工学报, 2012, 33(10): 124-126.
- HAN Xiao-hai, ZHANG Yao-hui, SUN Fu-jun, et al. Index weight determination method based on principal component analysis[J]. Journal of Military Engineering of Sichuan, 2012, 33(10): 124-126.
- [10] 车国鹏, 刘永红. 遗传算法优化 BP 神经网络的交通流参数预测[J]. 综合运输, 2018, 40(6): 64-67, 108.
- CHE Guo-peng, LIU Yong-hong. Traffic flow parameter prediction based on BP neural network optimized by genetic algorithm[J]. Integrated Transportation, 2018, 40(6): 64-67, 108.
- [11] 王杰, 卢毅. 基于基因表达式编程的地铁隧道土建工程造价估算[J]. 长沙理工大学学报: 自然科学版, 2019, 16(3): 17-24.
- WANG Jie, LU Yi. Cost estimation of subway tunnel civil engineering based on gene expression programming[J]. Journal of Changsha University of Science & Technology: Natural Science, 2019, 16(3): 17-24.
- [12] 邵雅楠. 基于主成分分析和 BP 网络的高层住宅工程造价估算方法[D]. 秦皇岛: 燕山大学, 2015.
- SHAO Ya-nan. Cost estimation method of high-rise residential project based on principal component analysis and BP network[D]. Qinhuangdao: Yanshan University, 2015.
-
- (上接第 67 页)
- [14] 李解, 王建平, 许娜, 等. 基于文本挖掘的地铁施工安全风险事故致险因素分析[J]. 隧道建设, 2017, 37(2): 160-166.
- LI Jie, WANG Jian-ping, XU Na, et al. Analysis of safety risk factors for metro construction based on text mining method[J]. Tunnel Construction, 2017, 37(2): 160-166.
- [15] Yu C T, Salton G. Precision weighting-an effective automatic indexing method [R]. Ithaca: Cornell University, 1975.
- [16] 石英, 孟玄喆. 基于轨迹交叉理论的制造业生产安全问题研究[J]. 工业工程与管理, 2014, 19(4): 129-134.
- SHI Ying, MENG Xuan-zhe. Analysis of safety risk factors for metro construction based on text mining method [J]. Industrial Engineering and Management, 2014, 19(4): 129-134.
- [17] Agrawal R, Imieliński T, Swami A. Mining association rules between sets of items in large databases[A]. Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data[C]. New York: ACM, 1993: 207-216.
- [18] Nenonen N. Analysing factors related to slipping, stumbling and falling accidents at work: application of data mining methods to Finnish occupational accidents and diseases statistics database[J]. Applied Ergonomics, 2013, 44(2): 215-224.
- [19] Agrawal R, Srikant R. Fast algorithms for mining association rules[A]. Proc. 20th Int. Conf. Very Large Data Bases(VLDB)[C]. San Francisco: Morgan Kaufmann Publishers, 1994: 487-499.
- [20] 郭圣煜, 骆汉宾, 滕哲, 等. 地铁施工工人不安全行为关联规则研究[J]. 中国安全生产科学技术, 2015, 11(10): 185-190.
- GUO Sheng-yu, LUO Han-bin, TENG Zhe, et al. Research on association rules of unsafe behavior for metro construction workers[J]. Journal of Safety Science and Technology, 2015, 11(10): 185-190.